CLAIMS

1. A method including

executing, on a computing device, an autonomous software element, said autonomous software element having access to information about a plurality of gene expression values;

generating, in response to said information, a hypothesized relationship about genes associated with said gene expression values, said hypothesized relationship having the properties of being (1) relatively unlikely to be due to chance, and (2) relatively likely to be of interest to at least one agent other than said autonomous software element;

sending information about said hypothesized relationship to said at least one agent.

2. A method as in claim 1, including collecting said information from a plurality of relatively nonlocal databases.

3. A method as in claim 1, including collecting said information from at least one relatively nonlocal database.

4. A method as in claim 1, wherein generating includes

selecting a first set of genes in response to said gene expression values;

1 selecting a second set of genes in response to data other than said gene ex-

2 pression values;

3 applying a statistical technique to said first set and said second set; and

4 confirming said hypothesized relationship in response to applying said sta-

5 tistical technique.

6

7 5. A method as in claim 1, wherein generating includes

8 selecting a set of genes in response to said gene expression values;

9 comparing the frequencies of the nucleotide sequences upstream from said

10 set of genes to the frequencies of the nucleotide sequences upstream of genes not in said

11 set; and

12 constructing a hypothesis that is responsive to sequences that have an

13 anomalous frequency distribution.

14

15 6. A method as in claim 1, wherein generating includes

16 examining pathways of genes with respect to gene activation sequences;

17 extending said pathway using said gene expression information; and

18 constructing a hypothesized relationship concerning an extension of said

19 pathway.

20

21 7. A method as in claim 1, wherein generating includes

1   evaluating correlation values with respect to an identified gene or gene se-

2 quence so as to determine variations in the behavior of said identified gene or gene se-

3 quence; and

4   confirming said hypothesized relationship in response to applying said sta-

5 tistical technique.

6

7   8.   A method as in claim 1, including

8   rating said hypothesized relationship with a measure of interest by said

9 agent; and

10   determining whether to send said information such that said determination

11 is responsive to said measure.

12

13   9.   A method as in claim 8, including

14   determining a threshold in said measure of interest; and

15   sending said information in a manner responsive to said threshold.

16

17   10.   A method as in claim 8, wherein said measure of interest is periodi-

18 cally reevaluated.

19

20   11.   A method as in claim 8, wherein said measure of interest is respon-

21 sive to whether said hypothesized relationship is relatively simple and relatively unlikely

22 to be due to chance, or whether said hypothesized relationship relates to at least one of:

1 genes one or more researchers have indicated they are interested in, genes for which there

2 are published papers, selected domain-specific knowledge about gene expression.

3

4   12.   A method as in claim 1, wherein said autonomous software element

5 has access to collateral information other than gene expression values.

6

7   13.   A method as in claim 12, wherein said collateral information in-

8 cludes at least one of: information about interests of said at least one agent, information

9 about published papers, information about researchers, information about relationships

10 between genes.

11

12   14.   Apparatus including

13   memory recording information about a plurality of gene expression values;

14   an autonomous software element disposed on a computing device, said

15 autonomous software element having access to said memory;

16   said memory recording information about a hypothesized relationship be-

17 tween said gene expression values, said hypothesized relationship having the properties

18 of being (1) relatively unlikely to be due to chance, and (2) relatively likely to be of inter-

19 est to at least one agent other than said autonomous software element;

20   a communication link coupled to said memory and capable of sending in-

21 formation about said hypothesized relationship to said at least one agent.

22

1   15.   Apparatus as in claim 14, wherein said communication link is capa-

2   ble of collecting said information from a plurality of relatively nonlocal databases.

3

4   16.   Apparatus as in claim 14, wherein said communication link is capa-

5   ble of collecting said information from at least one relatively nonlocal databases.

6

7   17.   Apparatus as in claim 14, wherein

8   said hypothesized relationship includes (a) information about a first set of

9   genes, said first set of genes having been selected in response to said gene expression

10   values, and (b) information about a second set of genes, said second set of genes having

11   been selected in response to data other than said gene expression values;

12   said hypothesized relationship has been confirmed in response to a statisti-

13   cal technique applied to said first set and said second set.

14

15   18.   An apparatus in claim 14, including

16   a means for selecting a first set of genes in response to said gene expression

17   values;

18   a means for selecting a second set of genes in response to data other than

19   said gene expression values;

20   a means for applying a statistical technique to said first set and said second

21   set; and

1   a means for confirming said hypothesized relationship in response to ap-

2   plying said statistical technique.

3

4   19.   An apparatus as in claim 14, including

5   a means for selecting a set of genes in response to said gene expression val-

6   ues;

7   a means for comparing the frequencies of the nucleotide sequences up-

8   stream from said set of genes to the frequencies of the nucleotide sequences upstream of

9   genes not in said set; and

10   a means for constructing a hypothesis that is responsive to sequences that

11 have an anomalous frequency distribution.

12

13   20.   An apparatus as in claim 14, including

14   a means for examining pathways of genes with respect to gene activation se-

15 quences;

16   a means for extending said pathway using said gene expression informa-

17 tion; and

18   a means for constructing a hypothesized relationship concerning an exten-

19 sion of said pathway.

20

21   21.   An apparatus as in claim 14, including

1   a means for evaluating correlation values with respect to an identified gene

2 or gene sequence so as to determine variations in the behavior of said identified gene or

3 gene sequence; and

4   a means for confirming said hypothesized relationship in response to ap-

5 plying said statistical technique.

6

7   22.   Apparatus as in claim 14, said memory including

8   information associating said hypothesized relationship with a measure of

9 interest by said agent; and

10   a software comparator coupled to said said measure of interest and to a se-

11 lected threshold.

12

13   23.   Apparatus as in claim 14, wherein said autonomous software ele-

14 ment has access to collateral information other than gene expression values.

15

16   24.   A memory recording information including instructions, said in-

17 structions interpretable by a computing device, said instructions including

18   an autonomous software element having access to information about a plu-

19 rality of gene expression values;

20   a first software element coupled to said information and capable of gener-

21 ating a hypothesized relationship between said gene expression values, said hypothesized

22 relationship having the properties of being (1) relatively unlikely to be due to chance, and

1 (2) relatively likely to be of interest to at least one agent other than said autonomous

2 software element;

3 a second software element coupled to information about said hypothesized

4 relationship and capable of sending information about said hypothesized relationship to

5 said at least one agent.

6

7 25. A memory as in claim 24, including

8 information about a first set of genes selected in response to said gene ex-

9 pression values;

10 information about a second set of genes selected in response to data other

11 than said gene expression values;

12 information about said hypothesized relationship selected in response to

13 applying a statistical technique to said first set and said second set.

14

15 26. A memory as in claim 24, wherein generating includes

16 information about selecting a first set of genes in response to said gene ex-

17 pression values;

18 information about selecting a second set of genes in response to data other

19 than said gene expression values;

20 information about applying a statistical technique to said first set and said

21 second set; and

1   confirming said hypothesized relationship in response to applying said sta-

2   tistical technique.

3

4   27.   A memory as in claim 24,

5   information about selecting a set of genes in response to said gene expres-

6   sion values;

7   information about comparing the frequencies of the nucleotide sequences

8   upstream from said set of genes to the frequencies of the nucleotide sequences upstream

9   of genes not in said set; and

10   information about constructing a hypothesis that is responsive to sequences

11   that have an anomalous frequency distribution.

12

13   28.   A memory as in claim 24, including

14   information about examining pathways of genes with respect to gene acti-

15   vation sequences;

16   information about extending said pathway using said gene expression in-

17   formation; and

18   information about constructing a hypothesized relationship concerning an

19   extension of said pathway.

20

21   29.   A memory as in claim 24,

1   evaluating correlation values with respect to an identified gene or gene se-

2   quence so as to determine variations in the behavior of said identified gene or gene se-

3   quence; and

4   confirming said hypothesized relationship in response to applying said sta-

5   tistical technique.

6

7   30.   A method including

8   executing, on a computing device, an autonomous software element, said

9   autonomous software element having access to information about concerning a set of

10   genes and proteins;

11   generating, in response to said information, a hypothesized relationship in-

12   volving at least one of the following: a set of proteins, SNPs or chemicals pertaining to

13   said information;

14   sending information about said hypothesized relationship to said at least

15   one agent.

16

17   31.   A method as in claim 30, including collecting said information from

18   a plurality of relatively nonlocal databases.

19

20   32.   A method as in claim 30, including collecting said information from

21   at least one relatively nonlocal database.

22

1   33.   A method as in claim 30, including

2   rating said hypothesized relationship with a measure of interest by said

3   agent; and

4   determining whether to send said information such that said determination

5   is responsive to said measure.

6

7   34.   A method as in 30, including

8   determining a threshold in said measure of interest; and

9   sending said information in a manner responsive to said threshold.

10

11   35.   A method as in claim 34, wherein said measure of interest is periodi-

12   cally reevaluated.

13

14   36.   A method as in claim 34, wherein said autonomous software element

15   has access to collateral information.

16

17   37.   A method as in claim 34, wherein said collateral information in-

18   cludes at least one of: information about interests of said at least one agent, information

19   about published papers, information about researchers, information about relationships

20   between genes.